电子资源管理过程中数据质量控制研究
——以清华大学图书馆基于 ALMA 系统的实践为例*

武丽娜　窦天芳　周虹　刘阳
清华大学图书馆 北京 100084

摘要:[**目的/意义**]介绍清华大学图书馆借助 ALMA 系统对电子资源管理和质量控制的经验，可为同行图书馆在下一代图书馆服务平台的应用实践和电子资源数据质量控制方法提供参考。[**方法/过程**]采用文献调研法提出电子资源数据质量控制的必要性；采用问卷调查法对同行图书馆电子资源管理模式和困境进行调研，提炼共性问题。基于 ALMA 系统中对电子资源管理的数据模型、工作流程和管理模式，从电子资源库和单册层级、书目层级、发现系统和中央发现索引层级，提出相应的数据质量控制方法，为电子资源管理过程中数据质量控制提供实践指导。[**结果/结论**]电子资源管理是一个复杂的、跨越多平台的系统工程，图书馆应制定多层次、立体化的数据质量控制方案，提升电子资源管理和服务水平。

关键词：图书馆 电子资源 数据 质量控制

分类号：G250

1 电子资源管理过程中数据质量控制的必要性

1.1 电子资源数据质量控制相关标准

资源检索与获取是电子资源管理过程中数据质量控制水平的直接体现，也是图书馆服务评价相关标准中的重要指标。多个国际和国内标准，如 NISO Z39.7-2002（美国图书馆统计标准）、ISO 11620：2008（信息与文献 国际图书馆绩效指标）、ISO 16439:2014（信息与文献 图书馆影响力评估的方法与流程）、GB/T 29182-2012（图书馆绩效指标）中均包括资源发现获取指标[1]。国内相关研究也将资源存取便利性作为评估图书馆管理服务能力的一项指标[2]。美国国会图书馆发布的《2022-2026 年数字馆藏发展战略》提出应提升资源的可用性和可访问性、增强内容的可发现性，构建数字馆藏集成仓储发现系统[3]。下一代图书馆服务平台使得图书馆电子资源全流程管理成为可能，但也给业务管理和数据质量提出了更高的要求。系统迁移带来的部门和业务重组、电子资源馆员工作职责的重新定义，更使得电子资源管理的难点由订购等业务流程管理转移到资源发现和获取服务。

北美地区对电子资源管理的研究起步较早，如美国数字化图书馆联盟（digital library federation,DLF）在 2004 年发布电子资源管理报告[4]。此后，欧美地区学术图书馆与出版行业紧密合作，建立起一系列方便电子资源管理、揭示与传播的技术规范，并迅速在图书馆建立起对应的工作流程及工作规范[5]。ALMA 是在这些工作流程及规范的基础上开发的全媒体资源管理平台。因此，研究与分析 ALMA 电子资源管理过程中的数据质量控制，对学习同行电子资源管理的先进经验、提高国际出版商与图书馆之间数据共享交换水平、提升图书馆资源发现获取水平大有帮助。

电子资源经费占比高、数量大、类型多样、来源多、分布广的特点决定了其是图书馆资源管理的难点。但是，国内图书馆基于下一代服务平台对电子资源管理的经验普遍不足，资源数量和服务体量也存在不均衡、不匹配等情况。电子资源数据的质量受数据库商数据服务、电子资源馆员工作技能、系统功能支持等多重因素制约，因此其数据质量控制方法和规范成为图书馆及出版行业共同面对并亟待解决的问题。图书馆有必要建立电子资源管理过程中的数据质量控制方案，以提升图书馆对电子资源的管理和服务能力。

1.2 相关研究与实践现状

** 作者简介：武丽娜，资源建设部副主任，馆员，硕士，E-mail：wulina@mail.tsighua.edu.cn；窦天芳，副馆长，副研究馆员，硕士；周虹，副研究馆员，硕士；刘阳，助理馆员。

笔者采用文献调研法，对电子资源数据质量控制相关文献进行调研。2022年11月28-29日，在Web of Science核心合集用主题词【TS=(librar*) AND TS=(electronic resource*) AND TS=(data quality)】检索得到212个结果，去除无关文献后得到与电子资源数据质量控制有关文献23篇。在中国知网用主题词“图书馆”、“电子资源”和“数据质量”进行检索，得到12个结果。浏览这些文章标题和摘要，分析得出，国外图书馆引入下一代服务平台的时间较早，重点关注馆藏电子资源数据质量评估、馆藏管理模式、书目数据质量、期刊质量标准等方面，并认为电子资源数据质量、是否进行了揭示对发现服务有较大影响[6-8]，有必要进行相应的质量控制。国内相关文献报道较少，主要集中在ALMA系统建设[9-10]、平台对比研究[11-12]、电子资源管理[13-14]、电子资源的编目数据[15-17]、使用统计数据[18]等层面。电子资源管理过程中数据质量控制方面，仅见关于电子资源访问故障相关的报道[19-20]。

笔者采用问卷调查法，并通过邮件、微信、QQ等方式，对12个图书馆进行问卷调研。这12个图书馆包含了国内的6个ALMA用户和6个非ALMA用户。6个非ALMA用户（北京大学图书馆、北京航空航天大学图书馆、吉林大学图书馆、中南大学图书馆、西安交通大学图书馆、浙江大学图书馆）均为国内“双一流”高校，在电子资源管理方面有丰富的经验。6个ALMA用户（北京师范大学图书馆、香港中文大学深圳图书馆、澳门科技大学图书馆、南方科技大学图书馆、中科院高能所图书馆、东北师范大学图书馆）分布在不同地理区域、且系统上线时间不同，具有一定代表性。调查问卷内容涉及图书馆使用的管理系统、编目的电子资源类型、编目标准、电子资源数量、供应商提供书目数据的比例、书目数据获取途径、电子资源揭示馆员所在部门及分工、电子资源揭示中遇到的困难、读者反馈与电子资源后台数据揭示、资源发现相关的问题等。由于部分馆对相关工作开展不多，所以收集的问卷项目不是很完整。对调研结果汇总分析，得出如下几点结论：①这些样本覆盖的资源管理系统包括ALMA、汇文、ALEPH等，资源发现系统为PRIMO、SUMMON、超星发现等。②其中7个图书馆对电子资源进行编目，其他图书馆（如采用汇文系统的图书馆）不支持电子资源编目，或图书馆无专职馆员对电子资源进行编目。整体上看，图书馆电子资源揭示人力有限或未设置相应岗位。③电子资源的编目范围主要集中在电子图书、电子期刊、学位论文，以买断资源为主，尚未拓展到其他资源类型。④虽然电子资源有相关的最新标准（GB/T 3792-2021信息与文献资源描述），但国内图书馆尚未采用统一标准。⑤电子资源编目馆员通常与纸本编目业务在同一个编目工作组，其中部分原因是人力有限。⑥遇到的电子资源揭示和质量控制方面的困难包括：电子资源数据验收机制不完备、部分数据库商不提供书目数据、元数据质量差、中文资源揭示困难、读者在发现平台查找不到所需文章或全文访问获取故障等。

笔者以清华大学图书馆作为案例进行研究，主要基于两点考虑：①清华大学图书馆较早开展电子资源采购，经过多年建设积累了较大体量的电子资源，目前有938个数据库、845万册电子图书、17.4万种电子期刊和多种其他类型电子资源，这些资源分布在全球范围内200多个异构平台上。②清华大学图书馆为国内较早引进ALMA进行电子资源管理的图书馆，积累了丰富的电子资源管理经验。迁移到ALMA之前，数据库、电子期刊、电子图书分别由Metalib、SFX和OPAC进行管理，基于数据库导航、期刊导航、馆藏目录检索和获取。迁移到ALMA和PRIMO后，实现了这些资源的统一管理和发现。而发现平台可检索的资源除了馆员在后台管理的期刊、图书级数据外，还包括部署在云端的中央发现索引覆盖的期刊全文、图书章节、研究数据集等4亿多条元数据以及PRIMO本地资源，因此数据管理层级增加，数据质量控制难度加大。由此可见，清华大学图书馆的电子资源体量大、在国内引入ALMA时间较早，积累的电子资源数据质量控制经验较为丰富，可为同行对电子资源的管理及数据质量控制提供参考。

2 ALMA系统的电子资源管理数据模型和工作流程

ALMA系统实现了对电子资源从订购到揭示的全流程管理。PRIMO实现了面向用户的资源一站式检索与发现。在ALMA中，各类资源都有统一的三层结构数据模型（如图1所示），以

便实现资源的统一管理。对于同一个知识实体，比如电子期刊 *the library*，通常情况下，ALMA 里只有 1 条书目记录（metadata management system,MMS），其单册(portfolio)有 2 个，对应不同的可访问年限（portfolio coverage）。同时，这个知识实体所在的两个资源库（electronic collection）也有相应的书目数据。

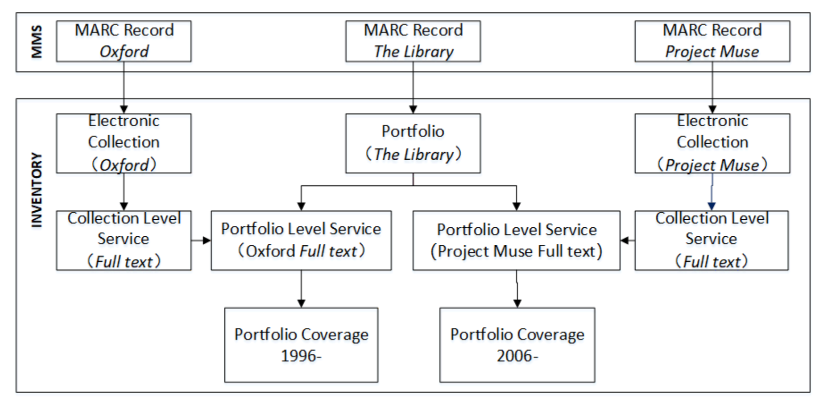


图 1 ALMA 中电子资源管理的数据模型

在 ALMA 中完成编目的资源会呈现在发现平台（PRIMO 系统界面）、数据库导航和电子资源导航等多个用户的使用场景中。ALMA 可作为链接解析器，与云端的资源元数据仓储实现从引文到全文的链接。作为 PRIMO 重要的数据模块，中央发现索引（central discovery index，CDI）基于 ALMA 的资源链接配置，对馆藏电子资源进行检索、发现及全文获取。ALMA 定时将图书馆的电子资源馆藏文件（holding file，HF）发布给 CDI。对期刊来说，其 ISSN、可获取年限范围等信息会被用于匹配 CDI 文章的来源期刊和年卷期等信息，并在发现平台显示是否有在线全文。因此，用户在发现平台检索一篇文章，从触发检索到定位到所在数据库，再到顺利获取全文，要经过多层级的数据匹配和映射。相应术语概念说明如表 1 所示，数据从 ALMA 到 PRIMO 的发布流程如图 2 所示：

表 1 ALMA 系统中电子资源相关基本概念

术语	**说明或示例**
馆藏（inventory）	某电子期刊及其所在资源库的链接参数、可访问年限等信息
电子资源库（electronic collection）	某数据库对应的资源包，如 Oxford University press journals
单册（portfolio）	电子期刊的可访问范围、链接参数等信息
服务（service）	电子资源库的服务包括全文类型及访问参数
书目（MMS）	电子期刊 *The Library* 的书目
中央发现索引（CDI）	Oxford 数据库 CDI 资源库：Oxford Journals 2020 social sciences
馆藏文件（HF）	ALMA 发布到 CDI 的馆藏文件，包含资源标识号、卷期信息等
CDI institutional profile	某电子资源库的 CDI 配置文件
链接解析器（link resolver）	ALMA 作为链接解析器接收并解析 OpenURL 信息，构建数字对象，并与书目数据中的元数据匹配，生成服务链接
共享区知识库（community zone，CZ）	ALMA 知识库提供的某数据库对应的电子资源库、单册、书目等信息

注：表中部分术语说明或示例来自图 1。

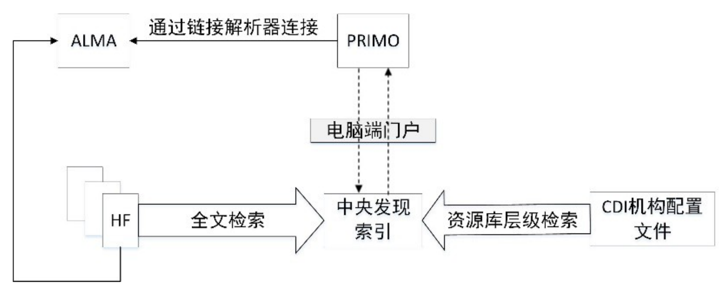


图 2　数据从 ALMA 到 PRIMO 的发布流程

注：参照 2019 年 CCEU 年会中国 Discovery Solution 进展报告（李珍,康利芬）中用图修订而成。

3 ALMA 系统对电子资源的管理模式

目前，对电子资源数据的揭示维护绝大部分依赖共享区知识库，知识库没有所需资源或资源覆盖不全、数据质量较差时，馆员需要对外源数据质量进行判断，用数据批处理软件核查或完善数据，然后通过导入书目数据、套录外部数据源等方式完成电子资源激活，如图 3 所示。具体采用哪种数据管理方式，主要取决于拟揭示资源的数据质量，包括元数据覆盖资源范围是否全面、字段是否完备、编目规则等。根据经验，外文出版商的数据质量要高于集成商；与进出口公司和数据公司有合作模式的数据库商能提供更高质量的元数据；而部分中文数据库商提供电子图书书目数据的能力和服务方面则有所欠缺，需要提升数据服务意识。例如，个别出版商会面向图书馆开展书目数据质量调查，调查内容涉及资源管理系统的选择、书目数据的获取、处理方式和工具等，并关注元数据与资源发现的关系[21]。图书馆通过外国教材中心联合订购的教材电子书、通过易阅通平台订购的电子书，由于进出口公司有相应的书目数据编目员，往往可以提供较高质量的书目数据，并能根据图书馆要求完善书目数据。而中文电子书平台、特别是古籍类、音视频类等资源类型通常只能提供 Excel 形式的元数据，需要馆员自行编制书目数据。图 4 是馆员利用 Excel、Marcedit 等工具完成批量编目的书目数据样例。

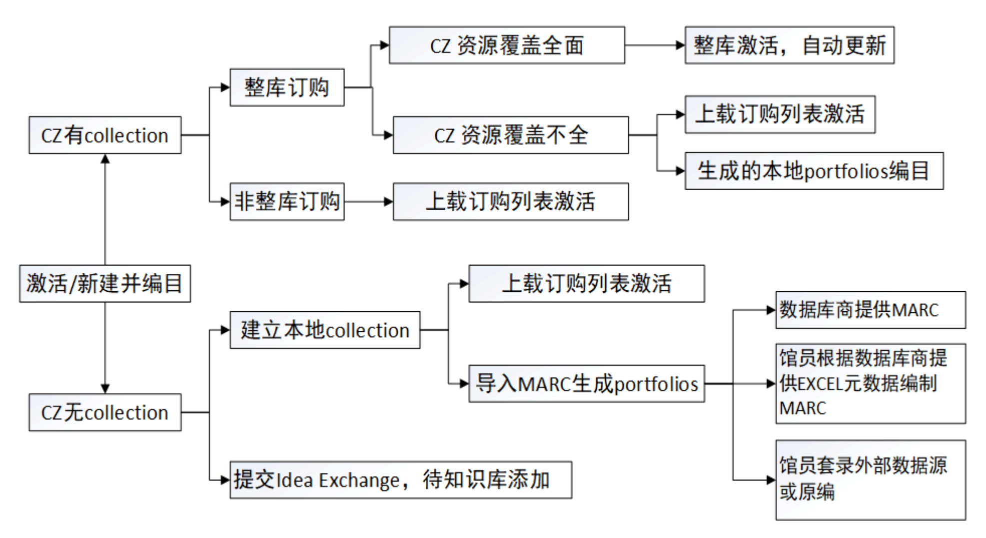


图 3 ALMA 中对电子资源的管理方式

```
LDR 01228nam0a22002771 450
001 991021699837203966
005 20200109154010.0
100 __ |a 20191228d1935 em y0chiy50 ea
101 0_ |a chi
105 __ |a y z 001zy
135 __ |a dr
2001_ |a 丘隅意见 |c 北牕琐语 |c 谈辂 |c 笔麈 |c 二酉委谭摘录 |f (明) 乔世宁撰 |f (明) 余永麟著 |f (明) 张凤翼撰 |f (明) 莫是龙撰 |f (明) 王世懋撰
210 __ |c 商务印书馆 |d 1935
215 __ |a 在线资源(151 页)
301 __ |a 文件号: XpoplUxU
330 __ |a 丘隅意见一卷，据百陵学山本影印。北牕琐语一卷，据砚云甲乙编本影印。谈辂一卷，据百陵学山本影印。笔麈一卷，据奇晋斋丛书本影印。二酉委谭摘录一卷，据纪录汇编本影印。
423 _0 |1 2001 |a 北牕琐语 |7 01 0 |a 余永麟, |f 明代
423 _0 |1 2001 |a 笔麈 |7 01 0 |a 莫是龙, |f 明代
423 _0 |1 2001 |a 二酉委谭摘录 |7 01 0 |a 王世懋, |f 明代
423 _0 |1 2001 |a 谈辂 |7 01 0 |a 张凤翼, |f 明代
610 0_ |a 中国丛书
701 _0 |a 莫是龙 |4 撰
701 _0 |a 乔世宁 |4 撰
701 _0 |a 王世懋 |4 撰
701 _0 |a 余永麟 |4 著
701 _0 |a 张凤翼 |4 撰
801 _0 |a CN |b TSU |c 20191228
856 40 |u http://www.hwshu.com/front/singleBookDetail/p/index/12933.do
```

图 4 利用 Excel、Marcedit 等工具完成批量编目的 MARC 数据样例

4 各数据层级的数据质量控制方案

基于上述 ALMA 系统电子资源管理的数据模型和工作流程，以及 ALMA 电子资源管理方式的介绍可发现，电子资源馆员不仅需要理解数据从 ALMA 到 PRIMO 的发布流程，也应熟悉电子资源管理的具体操作方法。本部分将结合电子资源管理数据模型和流程中涉及的关键数据层级，包括电子资源库和单册层级、书目层级、发现系统和中央发现索引层级，分别提出相应的数据质量控制方法，以提高资源的发现获取水平。整体来看，电子资源库、单册和书目层级的数据质量既依赖于数据库商提升数据服务和共享意识，也需要专业馆员提升对数据质量及其对用户检索影响的认识，并掌握电子资源流程管理和数据质控方法。而发现系统和中央发现索引层级数据质量控制更多依赖于数据库商对数据的规范以及系统厂商对系统功能的完善。

4.1 电子资源库和单册层级的数据质量控制

对于电子资源库（collection），共享区资源库会发生资源库删除、合并等更新，这些变更会导致激活状态改变，需要馆员及时追踪处理。对于单册（portfolio）层级的数据质量控制，馆员需要关注单册的资源类型、访问参数、书目数据等信息，并据此制定相应的处理方案。就清华大学图书馆而言，为了提升数据核查的主动性和维护的及时性，馆员在统计端配置了相应的统计报告，定期读取和了解知识库更新带来的变化，并据此进行周期性维护。知识库数据更新提升了资源数据准确性，为用户在 PRIMO 端检索数据提供便利，但也带来了馆员维护核查工作量增加等问题。知识库单册数据的更新包括资源可访问年限信息更新、资源删除、资源书目数据完善等方面。比如，电子图书书目数据更新后补充了 505 字段图书章节信息，更为完善。

除了关注知识库数据的变化，在实践中还采用了多种方式，如建议将数据验收环节纳入采访员工作职责；如对支持 KBART 数据收割的 Springer、Taylor 等数据库配置自动更新；如

对电子书数据进行编目时尽量以馆内纸本书书目数据为基础数据，提升数据质量并便于发现获取。

4.2 书目层级的数据质量控制

电子资源经过著录、标引成为书目数据，才能为用户提供更丰富的检索点。资源发现平台可实现对资源类型、作者、出版等信息的分面检索。而检索结果的准确率、查全率取决于 ALMA 中的书目数据质量。清华大学图书馆较早开展电子资源批量编目，认为国内图书馆整体对数据质量的重视程度不够[15-16]，批量编目是促进电子资源获取的有效手段。随着数据库商数据服务意识的提高，更多数据库商开始主动关注书目数据质量对资源发现的影响。

对于外文图书、外文期刊等，部分书目数据可依赖共享区知识库的数据，部分则依赖数据库商提供的书目数据或资源列表。但共享区知识库书目数据也存在题名著录不准确、缺少主题、作者字段等情况，质量良莠不齐。有多个 ALMA 用户建议知识库对其提供的书目数据进行评估，告知图书馆书目数据编目等级、ISBN/ISSN、题名、著者、主题字段的情况等，方便图书馆判断[22]。对于中文图书、中文期刊、视频等资源，存在着共享区知识库无书目数据、而数据库商提供书目数据不规范、缺字段、乱码或无法提供书目数据等情况，需要馆员修订或完全舍弃知识库书目数据。中文电子书数据源质量不高通常也体现在其自身平台数据不规范的情况[23]。有同行比较了用户通过发现平台查找和获取流媒体资源的情况，评估了流媒体资源的书目数据，并认为元数据质量对发现服务有较大影响[24]。而笔者在实践中也发现，平台书目数据质量不高时，数据库商往往也难以提供高质量的书目数据。表 2 是清华大学图书馆电子图书书目数据基本情况以及采取的不同数据处理方式，针对不同数据库数据源特别采用不同的数据处理方式，保障了电子图书数据的揭示更为全面准确。

表 2 电子图书数据库书目数据基本情况

类型	书目记录				典型数据库	数据处理方式
	记录来源/获取方式	格式	规则	载体形式		
中文	数据库商提供（数据库商主动提供/馆员索要/图书馆倡议数据库商与数据公司合作后提供）	CAMARC	CALIS/国图/CCBD 等	纸本	科学文库、橙艺数据库	纸本书目记录转换成电子书目记录后导入
		Excel	—	—	瀚文民国图书、党政图书馆	馆员将 Excel 中元数据批处理成 MARC 数据后导入
		XML	—	电子	超星电子图书	XML 直接导入，根据 035 字段匹配增删数据
外文	管理员平台下载	MARC21	AACR2/RDA	电子	Academic Video Online、Elsevier	直接导入 ALMA
	数据库商提供				PQDT、World Library	导入 PRIMO 本地

对电子资源的书目数据进行核查和批处理是电子资源数据质量控制的重要一环，对馆员的胜任力有较高要求。某大学图书馆对编目和元数据助理馆员的任职资格中提到，馆员除需

具备编目工作经历外，还需具备 ALMA 系统书目数据导入和批处理能力，对不同类型资源应用不同元数据标准的能力，理解发现平台元数据相关问题的能力、对关联数据的理解能力等[25]。在对电子资源数据批处理的实践中，馆员还应掌握数据批处理软件 Marcedit、Excel 等工具的使用、ALMA 中数据导入配置等的使用。因此，馆员在实际工作中应主动加强掌握各种数据批处理方法，提高书目数据批处理效率和数据质量。

4.3 发现系统和中央发现索引层级的数据质量控制

ALMA 系统的书目记录会根据一定规则（normalization rules）转录为相应 XML 文件后入库，用于读者查询和展示。虽然发现系统（PRIMO）并不会对来自 ALMA 系统的元数据本身进行修改，但是可以根据一定的规则来限制元数据中的信息是否能被检索或显示。PRIMO 系统对数据质量的控制主要来自规则编制。这套规则的编制方式非常灵活，可根据使用者的需求，对不同类型资源的元数据展示进行按需定制。但由于元数据转入时，图书馆可以在展示、检索、排序、去重、分面等方面进行订制，这些设置最终会影响书目记录在用户端的展示和检索效果。因此，图书馆需不断调整使规则编制实现最优化。

实践中，馆员可以根据需要设定条件，来决定发现系统中不同类型资源的呈现内容；也可以根据去重算法，编辑 ALMA 系统中的记录，PRIMO 系统会依据识别符等信息将这些数据进行匹配合并操作，形成一条更完整的逻辑记录，使同种资源的不同类型集中在一条书目记录中呈现，方便用户获取资源。但在进行这些有意识的通过系统规则影响记录呈现的时候，其实更应该考虑用户是否需要这种合并或者去重的便利。在实际使用中，有些图书馆对系统进行限制，不再合并同种资源不同类型的记录，而是选择将所有记录原样呈现，供用户自己进行挑选。同时，我们也应该看到，用户的检索习惯受网络搜索引擎的影响，输入的检索条件越来越宽泛，一定程度上需要系统去根据一些用户行为习惯来猜测用户实际想要的结果。而发现系统现在要兼顾查全率和查准率，仅从记录的查重及合并上来做，已经达不到要求，而是需要从发现系统本身的搜索引擎上来改进，要做到这一点需要厂商更多的技术支持，图书馆本身已经很难实现。

中央发现索引（CDI）的元数据来自 1 600 余个出版商、内容提供商、机构库中超过 40 亿条的元数据，包括期刊文章、图书篇章记录、数据集合等。针对 CDI 层级的数据质量控制，国外同行曾从内容、检索功能、发现、交互性和易用性 5 个方面的 36 个指标对用户对资源发现平台的使用进行调查分析，用户认为易用性最重要，中央索引数据质量权重最高[26]。笔者及同行都发表了关于下一代服务平台环境下电子资源访问问题与影响因素相关的文章，认为 ALMA 知识库数据质量、元数据质量、电子资源自身信息变更等是电子资源访问故障的主要原因[13,20]。对于某个电子资源库来说，其文章级检索与获取除了与单册数据、书目数据、资源从 ALMA 到 PRIMO 的发布有关外，与资源库的 CDI 数据配置也有直接关系。电子资源馆员一方面需要对电子资源库的 CDI 菜单进行配置和维护，另一方面应了解相应数据库的篇章级资源元数据的覆盖范围。也可根据该资源库的 CDI 数据质量，决定是否支持篇章级资源在发现端的显示。这些都需要馆员对各数据库的数据情况及系统配置进行深入的测试分析，据此制定配置方案。

5. 数据质量控制的建议和举措

除上述质量控制外，图书馆还可在加强文献资产和文档制度规范管理、充分利用系统功能实现数据自动收割和对接、基于统计分析进行质量控制、建立与师生间有效的沟通机制和渠道、加强馆员能力培养和团队协作、加强数据共建共享六个方面开展工作。

5.1 加强文献资产和文档制度规范管理

文献资产管理是图书馆服务的基础。图书馆应构建精准化的文献元数据资产管理体系，

改变其“有资源无数据”的窘境，使图书馆成为真正意义上资源的所有者[27]。笔者认为，应对数据库商提供的资源列表提出明确要求，并建议将数据全面性、准确性要求加入资源订购合同。此外，图书馆应对来自数据商的订购资源数据列表、书目数据进行有效存档。电子资源管理头绪多，需要有相关的文档制度进行规范管理。比如，建立适合本馆的电子期刊揭示流程和规范[28]，基于各个数据库建立馆藏数据维护手册、做好元数据规范存档，这些工作应该贯穿于电子资源管理的整个过程。

5.2 充分利用系统功能实现数据自动收割和对接

对于系统支持自动收割的数据库，如 Springer、Elsevier、Wiley 等，图书馆可与数据库商沟通获取数据收割参数进行配置，实现资源按周期自动收割。对于可从平台下载订购数据的库，如 Springer，馆员可自行从平台下载并导入 ALMA 系统支持的 KBART 格式。这两种方式都可以提高数据准确性，同时减轻馆员维护量。资源使用数据，如 COUNTER 报告，也建议配置自动收割协议，实现数据自动对接。

5.3 基于统计分析进行质量控制

有国外同行认为，ALMA 系统统计分析功能可帮助图书馆发现资源访问获取故障、了解业务趋势和用户使用行为[29]。虽然研究这些模块的数据对馆员来说是非常耗时的，但是图书馆有必要对馆员进行培训并使得更多馆员掌握建立数据报告的技能并成立团队进行分析数据管理。ALMA 统计端提供了多个访问服务统计功能。比如，根据题名检索排序的报告，可以判断馆藏发展应该关注哪些方面。通过使用链接解析器获取资源失效的报告，可以关注如何降低无服务的数据。在实践中，清华大学图书馆会用报告“Top ten article title accesses via OpenURL requests without services in the previous year”追踪访问故障原因并进行解决。对数据层面的深入分析可以以点带面地发现数据库层级的问题。这个模块的运用，不仅可以帮助馆员了解读者对资源的使用需求、遇到的困难，也提高了馆员的数据管理能力和服务意识。

5.4 建立与师生间有效、多渠道的反馈与沟通机制

馆员可以通过座谈交流、问卷调查、资源使用数据分析、用户使用行为分析等多种方式了解师生对电子资源的使用习惯和使用体验，使得图书馆对资源的管理与服务需求更好地融合。清华大学图书馆通过多个渠道接收师生关于电子资源获取的反馈，包括资源发现平台“问题反馈”表单、馆内电子资源微信交流群、学科馆员或馆际互借反馈的来自师生的需求和咨询等。馆内 PRIMO 系统界面水木搜索的“问题反馈”栏目，能够接收到师生关于资源荐购、访问与获取等方面的建议，其中与电子资源访问与获取方面的建议集中在平台数据、元数据错误、全文获取失效、校外访问、使用权限等方方面面，成为图书馆与用户沟通的桥梁。

5.5 加强馆员能力培养和团队协作

电子资源管理模式和系统的转变带来业务和机构重组，催生了图书馆新架构[30]。电子资源馆员核心能力培养也受到业界关注。2017 年，美国图书馆协会（american library association,ALA）下属的图书馆馆藏与技术服务协会理事会发布《编目和元数据专业馆员核心能力》标准[31]。北美连续出版物兴趣（north american serials interest group,NASIG）执行委员会于 2013 年批准采用的一项电子资源馆员职业能力标准包括七项核心能力[32]。有研究认为，目前国内电子资源馆员的核心能力水平整体偏低[33]。图情领域应在教学课程设置和在线培训两个方面更加关注图书馆员自身的职业发展。一方面，在图情领域课程设置上，应更加注重与实际工作的联系。另一方面，图书馆应借助美国图书馆协会在线平台[34]、图书馆员职业发展教育平台 Library Juice Academy[35]等为馆员提供能力提升机会。图书馆应思考有效的岗位配置和任务分配方式，加强电子资源管理知识体系培训，支持馆员技能提升。另外，电子资源管理需要更多

的团队协作，应建立相应的协作机制和流程，发挥采访馆员、揭示馆员、参考咨询馆员、技术馆员各自的优势，有效协调协作，提升应对和解决相关问题的能力，使得馆员在电子资源整体化管理模式中找到自己的位置。

5.6 加强与系统开发商和资源共享联盟的建设与合作

下一代服务平台环境驱动图书馆融入电子资源管理的整个生态圈，图书馆独立管理电子资源的难度加大，业界开始寻求以联盟的方式统一管理电子资源。一项关于图书馆联盟采用 ALMA 系统情况的调研显示，图书馆间以联盟形式选用 ALMA 系统进行统一资源管理和服务的联盟数量呈上涨趋势[36]。例如，日本国立情报学研究所引进 ALMA 系统为日本大学图书馆电子资源采购联盟 500 余家图书馆提供电子资源管理支持[37]。因此，图书馆一方面应进一步加强与系统开发商的沟通和深度合作，结合本地化需求进行开发升级，提升系统性能和国内服务能力；另一方面，应借助联盟支撑作用加强合作，借助 DRAA、CALIS、CASHL、NSTL 等机构力量以及馆际合作共享数据等方式，寻求与下一代图书馆服务平台的融合发展和数据共享。图书馆不仅要与系统平台商、联盟之间加强沟通合作，也应与出版行业联手探讨更科学、更高效的电子资源揭示方式，形成数据质量控制提升的合力。

6 结论

下一代图书馆服务平台的应用、电子资源的特点和用户需求的多样化,对图书馆乃至整个电子资源生态圈都提出了挑战，对图书馆及电子资源的管理和服务能力提出了更高要求。图书馆应因势而变，积极应对变化，制定多层次、立体化的电子资源数据质量控制方案，提升自身电子资源管理和服务水平。本文调研了 12 个同行图书馆在电子资源数据揭示和质量控制方面遇到的共性问题。并介绍了清华大学图书馆基于 ALMA 系统的多层级数据质量控制方案，还从制度建设、技术支撑、培训、合作等角度提出了多项提升电子资源数据质量的建议，期待这些实践经验可以为同行图书馆提供借鉴参考，并引起整个出版商、系统商等电子资源生态圈中各个主体对电子资源数据质量控制的关注。

参考文献

[1]崔竞烽,郑德俊.国内外图书馆服务评价指标演变[J].图书馆论坛，2022,42(4):1-11.

[2]叶继元,郭卫兵,郑德俊,等.高校图书馆质量评价指标体系框架探讨[J].中国图书馆学报,2021,47(252):53-66.

[3]Library of Congress Digital collections strategy overview 2022-2026[EB/OL].[2023-04-06]. https://www.loc.gov/acq/devpol/Digital Collections Strategy Overview_final.pdf?loclr=blogsig.

[4] Digital library federation.Electronic resource management:report of the DLF electronic resource management initiative[EB/OL].[2023-03-20].https://old.diglib.org/pubs/dlfpubs.htm.

[5] Association of college and research libraries.Standards for libraries in higher education[EB/OL].[2023-04-06].https://www.ala.org/acrl/standards/standardslibraries.

[6]DAVID V K, HIKARU N, GERALD L, et al. Managing bibliographic data quality for electronic resources[J].Cataloging & classification quarterly,2017,55(7/8):560-577.

[7]DHARMENDRA T,ATUL B,MAYANK T, et al. Assessment of e-service quality

performance of university libraries[J].Digital library perspectives,2021,37(4):384-400.
[8]NERO M D, HE J.Is it necessary: quality control in cataloging?[J].International
journal of librarianship,2018,3(2):85-95.
[9]窦天芳,杨慧.清华大学图书馆一体化资源管理平台建设—以ALMA系统实施为例[J].数字图书馆论坛,2020(5):2-7.
[10]刘斌.下一代图书馆服务平台Alma——以北京师范大学图书馆为例[J].图书情报工作,2019,63（4）：79-85.
[11]高斌.图书馆云服务平台的发展趋势——以ALMA与WMS系统为例[J].图书馆研究与工作,2020(10):27-32.
[12]王晓翠.下一代图书馆服务平台ALMA与FOLIO的对比分析[J].数字图书馆论坛,2020,(12):2-8.
[13]武丽娜,贾延霞,杨慧,等.电子期刊有效揭示和维护的实践与思考[J].图书情报工作,2018,62(12):46-50.
[14]田晓迪.基于Alma的电子资源管理变革及实践思考[J].图书馆杂志,2021,367(11)：64-71.
[15]贾延霞,杨慧.Excel和元数据处理工具在电子资源批量编目中的应用[J].图书馆杂志,2014,33(1):40-44.
[16]贾延霞,杨慧.电子资源批量编目的实践及研究[J].图书情报工作,2014,58(18):117-121.
[17]汪初芸,苏建华.图书馆电子资源编目研究进展分析[J].现代情报,2016,34(4):134-136.
[18]朱玲,崔海媛.高校图书馆电子资源使用监控与统计系统数据获取质量评估方法探讨[J].图书情报工作. 2016,60(5):51-57.
[19]武丽娜,贾延霞,窦天芳,等.图书馆新服务平台环境下电子资源访问故障研究及实践[J].图书情报工作，2020(6):20-26.
[20]李峰,黄婧.新一代图书馆服务平台下电子资源访问问题探析—基于AlMA与PRIMO系统[J].图书馆学研究，2020(13):49-55.
[21]Elsevier.MARC数据需求调查—Elsevier ScienceDirect 电子图书[EB/OL].[2023-04-06].https://forms.office.com/r/BHAX8ieS4w.
[22]ExLibris.Provide MARC record quality measures to CZ collections[EB/OL].[2023-04-06].
https://ideas.exlibrisgroup.com/forums/308173-alma/suggestions/39648964-provide-marc-record-quality-measures-to-cz-collect.
[23]詹丽华.我国中文电子图书书目数据揭示研究[J].图书馆杂志，2021(10):64-71.
[24]JIAN W, ELSA L.The library has infinite streaming content, but are users infinitely content? the library catalog vs. vendor platform discovery[J].Journal of electronic resources librarianship.2020,32(2):71-86.
[25]Indeed.Assisant/associate librarian,cataloging and metadata[EB/OL].[2023-04-05].https://www.indeed.com/jobs?q=Library
Cataloger*&start=10&vjk=1f527e6e823ddd23.
[26]KIM S.Analysis of relative importance on evaluation elements of library discovery[J]. Journal of the Korean library and information science,2020,54(2):399-417.
[27]赵立冰，高景山．图书馆电子资源资产管理的思考[J]. 内蒙古科技与经

济.2012,9(259):158-159.
[28]武丽娜,贾延霞,杨慧,等.电子期刊有效揭示和维护的实践与思考[J].图书情报工作.2018,62(12):46-50.
[29]SHARP C.Alma analytics: a useful tool for access services?[J].Journal of access services,2021,18(1):17-29.
[30]刘素清.电子资源管理催生图书馆新架构[J].大学图书馆学报,2014(2):18-23.
[31]American library association.Core competencies for cataloging and metadata professional librarians[EB/OL].[2023-04-06].https://alair.ala.org/bitstream/handle/11213/7853/Core Competencies Cataloging Metadata Professional.pdf?sequence=1&isAllowed=y.
[32]NASIG.Core competencies for electronic resources librarians [EB/OL].[2023-04-06].https://www.nasig.org/site_page.cfm?pk_association_webpage_menu=310&pk_association_webpage=7802.
[33]叶岚,赵雅,王英,等.我国电子资源馆员核心能力调查与分析[J].大学图书馆学报,2020(6):55-62.
[34]American library association.ALA eLearning[EB/OL].[2023-04-05].http://www.ala.org/educationcareers/elearning.
[35]Library Juice Academy. Oline Professional developmen for librarians. [EB/OL].[2023-04-06].http://www.libraryjuiceacademy.com/courses.php.
[36]GUOYING L,PING F.Shared next generation ILSs and academic library consortia: trends, opportunities and challenges[J]. International journal of librarianship,2018, 3(2):53-71.
[37]JUSTICE.News.[EB/OL].[2022-06-17].https://contents.nii.ac.jp/en/justice.

作者贡献说明：
武丽娜：确定论文选题，开展文献调研，撰写问卷调查、电子资源管理模式、书目层级质量控制等部分内容，文章统稿与修改；
窦天芳：论文指导与修改；
周虹：撰写发现系统数据质量控制部分内容；
刘阳：撰写电子资源库和单册质量控制部分内容。

Research on Data Quality Control in Electronic Resource Management: Taking the Practice of Tsinghua University Library Based on ALMA System as an Example

Wu Lina Dou Tianfang Zhouhong Liuyang
Tsinghua University Library, Beijing, 100084

Abstract: [**Purpose/Significance**] This paper puts forward the experience of Tsinghua University Library in electronic resource management and quality control based on ALMA system, which can provide reference for the application practice of peer libraries in the next generation library service platform and the method of electronic resource data quality control. [**Method/Process**] The necessity of data quality control of electronic resources was proposed through literature research; This paper adopted the method of questionnaire to investigate the electronic resource management mode and predicament of peer libraries, and extracted common problems. Based on the data model, workflow and management mode of electronic

resource management in ALMA system, this paper put forward corresponding data quality control methods from the electronic resource library and single volume level, bibliographic level, discovery system and central discovery index level, providing reference for data quality control in the process of electronic resource management. [**Result/Conclusion**] Electronic resource management is a complex and multi-platform system project. The library should develop multi-level and three-dimensional data quality control scheme, and improve the electronic resource management and service level.

Keywords: library electronic resources data quality control